下载APP

14 | 不够坚定：异地多活没有一步到位的遗憾

2021-11-10 利莹

《超级访谈：对话张雪峰》 课程介绍 >

讲述：吴骏龙、利莹

时长 13:14 大小 12.12M

极客时间：饿了么异地多活一次成功算是互联网技术圈挺浓墨重彩的一笔，当时你们决定做这件事之前，经历了一个什么样的过程，可以讲讲吗？

张雪峰：2017 年之前，整个公司所有的业务系统都是部署在北京机房，服务器大概有四千多台，另外还有灾备的机器是在云端，都是虚拟机，大概有三千多台。当时我们峰值的业务订单数量已接近千万级别，但是基本上北京机房（IDC）已无法再扩容了，也就是说空余的机架，没有办法添加新的服务器了，必须再建一个新机房，于是我们在上海建了一个新的机房。

领资料

做多活之前，其实这里面有一个失误（严格说不算失误，法不责众，国内几乎皆如此）。当然这个失误我认为大部分团队可能都会犯，就是先搞灾备，不敢一下子上异地多活，没

能一步到位。其实当时我也提过“我们是不是一步到位”这个问题，虽然外界感觉饿了么一步到位了，其实我们之前还做过一个灾备，花了不少钱。

极客时间：所以一开始先做了灾备，才做了异地多活对吧？

张雪峰：对，这是常规路径。中国绝大部分公司都是做了灾备后才想做多活，当然能不能做成另说，有的最终没做成，或不敢做，或者内部阻力很大。

美团当年就没做成，现在有没有彻底搞定异地多活我不太清楚。其实美团技术底蕴挺强的，因为跟着王兴早期创业的人不少，包括美团还曾做过美团云。只要做过云，哪怕这个云再烂，都储备了一批做基础设施挺强的人，这不是一般公司技术底蕴能比的。但是美团业务太强势，停几个月业务需求，或只能做些无关核心的边角料需求，不可能的。

饿了么好在技术这块拧成一股绳，基本就是我一个决断声音，说难听点也是一言堂，或者我跟团队即使吵得再凶，他们再挑战我，但最后出去都是一个声音。我跟他们说，你们当场挑战我，都没问题，我也无所谓面子，但是既然我们吵架吵出结果，你就不要再回去反悔，不允许反悔，你跟团队反悔、抱怨也是不允许的。

曾经有一个同学，当场同意，回去又开始磨磨唧唧，在群里乱发，这个不行那个不行。哪怕这件事事后证明错的（谁能在结果出来前证明这是错的？），在当时你就得错到底，证明它是个错也是有价值的。更何况，异地多活最终一次上线成功，之后更是挽救饿了么线上系统无数次。当初有异议的同学，也包括外部知道这事但并不看好饿了么技术团队能搞定的同学都无话可说，虽然最初下这个决断有一定赌博成分。

唯一一次美团跟我们做过的官方技术交流，就是我们做了异地多活后。大家知道，虽然我们两家商业上竞争白热化，但技术同学间还可以，大家不敌视，交流比较坦诚。

极客时间：灾备和多活区别是什么？

张雪峰：灾备什么意思？灾备是我有一套一模一样的系统，每次发布的时候两边都发一次，但是生产的流量或者实际的数据不经过灾备系统，只经过主系统，然后通过数据同步模块全部同步过来。但它有很大缺陷和风险，如果做到极致，全是 AI 控制或机器人控制（包括发布后各种校验、验证等），理论上精度会很高，但实际很多环节还是人来控制。

我们发布很多，一天上千次，多的时候超过 2000 次，还有依赖，比如，你发了 A 、A.1、A.1.1，这里面只要有一个漏发，可能你生产系统没问题，因为你要用流量灰度切换的，但是那边（指灾备 IDC）是没有灰度的，没有流量，今天忘了就忘了，运气好明天又发了这个就 OK。但万一来一个事故，流量一下子进来，反而把数据搞乱了，结果不应该发的补贴发了，这也就算了，赔点钱，但是如果数据错了，就大麻烦了。所以它没有经过真实流量考验。很多业务场景，并不是发布系统默认双发或多发就能解决的，还是需要大量人工验证或灰度流量验证才能确保万无一失。

做灾备，你要解决两个问题，备份和恢复，备份和恢复永远是一对的，不能切开。一般的企业更多在备份上花功夫，这只做了一半，不常态化做恢复演练的团队或公司，已经有太多沉痛教训了。而多活呢，既解决了备份，又解决了恢复问题，因为不需要恢复，数据始终在跑。换句话说就是只要确认是 Workable，可工作的，就可以了。

极客时间：所以在多活之前做灾备其实没必要是吗？

张雪峰：不是，这个要具体问题具体分析的。

当时我们为什么做灾备？因为直接上异地多活，我是有犹豫的。跟我很熟的人会经常质疑我，在公开的会上也会挑战我（鼓励 Challenge Up 也是饿了么技术文化中自由、开放的一种体现），说“峰哥，你当时不应该做灾备”。但我当时的理由是说，中农工建都做了，美团也做了，滴滴也做了，除了阿里系（淘宝、支付宝算一类），他们确实直接搞了异地多活。那么，除了阿里系，在全中国大家想做交易类系统第二只螃蟹吗（根据事后复盘，严格说这并不算第二只螃蟹，而是第一只网格化环境下的异地多活螃蟹）？我说我不敢冒这个险。

所以代价就是你做了一个用处不大的东西，但它也有一定用处，万一被勒索软件盯上了，我们还有一份数据备份，但是从现实意义来看就好像没有太大用处。这个其实也很难评价，就比如说你做各种形形色色监控软件，你投入很多资源搞中间件，应该很有作用吧？但只要没出事，大家可能看不出来，业务团队或 CEO 估计就更难以体会在这方面进行投入的价值了。

所以，在国内有一个很怪的现象，这确实跟发达国家不一样。在国内，不管什么行业，救火英雄永远得到最大赞扬，幕后英雄却无人问津，导致很多真正牛逼的人、防范于未然的人，反而委屈地走了。我在饿了么，尽可能去扭转这个局面，但也没有完全扭转过来。因

为你很难界定什么叫防患于未然、什么是因为基础设施投入而带来的巨大价值或避免的巨大风险。

在技术决策上，这是我的一次偏差，但后来还是做了异地多活，当然皓哥（左耳朵耗子）和建刚他们让我坚定了决心，我当时不是反对，而是有犹豫。

极客时间：你不敢下决心搞多活，具体的原因是什么？纠结或者犹豫的点具体是什么？

张雪峰：我吃不准中间件这个团队能不能承担并搞定（除淘宝外）这种规模（日均常态千万级订单）下的交易系统上去实现异地多活。更何况，异地多活几乎只有一次上线机会（最大风险不在上线后稳定性保障，而在多活节点间的数据错乱），不成功则成仁。

只不过后面我下了决心，我们也不用立什么军令状了，如果干砸，从我开始，核心团队在收拾完烂摊子后（这是 15 年夏天那次张雪峰快扛不住想引咎辞职的挣扎后，对责任的全新理解）全部下台。

做异地多活的话，业务要耽误三个月，接不了涉及核心系统的大需求或大变更，业务部门压力非常大，我要去说服。而且上线后如果出问题，那就是致命的问题（数据错乱），因为数据是分开的（多个机房）。我们原来就一份数据，宕机了可以恢复，至少数据不会乱。所以宕机不是最可怕的，就怕你连宕机的机会都没有（之前提到的安全问题类似），数据错乱或丢失，是最可怕的。

如果按我对自己的要求的话，这件事我认为算是一个失误，说得好听点叫偏差，应该一步到位搞异地多活。因为做灾备也投入很多钱，也有媒体说钱打水漂了。但这是事后诸葛亮，我当时也没这样的勇气，因为确实饿了么那时经不起任何大的波浪。虽然这是给自己找的理由，但我感觉如果要求苛刻一点，确实是一个不正确的决定。

决定做异地多活之后，当时找了毕玄、阿里云交流，但是后来发现不能直接用阿里那套（所以前面提到，不是国内第二只螃蟹，是真真切切第一只），因为阿里是适配全网应用，淘宝是没有网格概念的，三亚和哈尔滨是没什么本质区别的。但我们不行，我们是一个表面广域网实际却由无数局域网构成的特殊业务网格，而且我们网格之间是有重叠的，还比较复杂。所以后来决定只能自研。

即使这样，我们第一批上线的异地多活还是把物流落下了，这是后话。

极客时间：如果现在来选的话，直接就搞多活了对吗？

张雪峰：对，如果我回到当年，我就知道这个团队是能承担这个使命的，肯定直接搞多活，当时我也不是太有底气。如果直接搞多活，也能省好多钱了（指投入到腾讯云做灾备的钱）。

极客时间：异地多活当时做完之后，其实效果还是非常好的是吧？

张雪峰：是的，这件事我最为团队骄傲的，不是说异地多活后面帮我们挡过很多以前挡不过的灾难，而是我们有惊无险一次切换成功。因为一旦切换不成功就是灾难，打补丁都很难（数据错乱比宕机更可怕），核心团队散架（士气打击）、我直接下台（业务问责）几乎是必然的。

支付宝有一次严重事故，宁愿宕机不敢切换，就是因为不能出现一分钱差错，这个简直就像银行卡里面出问题一样，一分钱就致命，所以其实当时我不是很有信心能一次搞定，异地多活又是做得跟阿里不一样套路，如果真出事，毕玄、阿里云也帮不上，都是我们自研适配的架构，那一次，真的是大家连续几天，彻夜不眠。

打个不恰当比方，双十一或 618 出事，最多是稳定性层面故障，舆论压力大但不至于对公司产生致命影响；但如果异地多活切换出事，舆论初期或许感觉不到，但一旦短时间不能修复（数据错乱，时间越长错乱会以几何级数增长），对公司绝对是致命影响，CTO 中“O”的价值倒是体现出来了，超级负面价值，或许应了我之前说过的一句话：你可能连宕机的机会都不会再有了。

极客时间：你刚刚也说到第一批上线异地多活，把物流落下了，这个是什么事情？

张雪峰：后来我反思这也是我另外一个不够坚定的地方，也不能说我心软吧，就是不坚定。因为当时物流业务老大，也是联合创始人康嘉，还有技术负责人刘昊旻都犹豫了，问我能不能把他们放第二批？物流出问题，他们都得“跳楼”。

但这样就等于帮物流填了个临时坑，却为整个技术团队挖了个更大的坑。后来我们要兼容第二批（其实第二批就物流，只不过物流涉及系统太多）上异地多活的物流，在这期间，要随时保证物流那边的顺畅，就做了很多开关，然后我们中间件团队、架构师团队还有 PMO 团队，把杂七杂八这两套东西并线。如果是两套业务系统并线还好一点，只要切个数

据库就行了，但这两套异地多活很麻烦，有一堆的风险要考虑。等于我们维持了几个月的“大部分多活 + 物流非多活”并行，切换时又胆战心惊地担心物流切换出问题，还得切回去继续并行。

我们原来设想异地多活只能一次性切换，因为我们的业务是强耦合的，不像携程，携程机票、酒店关联度不大的，你要订机票 + 酒店，做个简单聚合就行了，但我们不一样，饿了么是用户下了单，商户接了单，物流就要送单，上下游其实是强耦合（高内聚）。

程序员可能会说，现实业务没你说的那么理想，该强耦合就强耦合，其实不是强耦合，另一个词叫高内聚，该内聚的时候你不要去追求什么微服务那些乱七八糟的东西，就应该高内聚，因为就是一个业务形态，业务才是最重要的判断耦合或内聚的依据。谁（调用方 / 消费方）也离不了谁（被调用方 / 服务方），你每次调用都涉及到它，干嘛非强扯开来？没太大好处，当然，可以分开发布算一个好处，但也仅是技术上的好处，不是业务或领域上的好处。

最后送大家一句话，也是我对当年多活改造迁就物流团队的一个教训总结，用英文描述比较贴切：Avoid solving one problem by creating a BIGGER one，希望大家今后都能顺利跨过这类坎，不再重蹈我当年覆辙。

极客时间：你刚刚说美团也想做异地多活是吗？相当于它也是走了常规路线，先做了灾备，再想做多活。

张雪峰：是，美团做了异地灾备。美团现在没有 CTO 了，在王慧文之前，CTO 叫罗道锋，他原来是大众点评的 CTO，后来美团和大众点评合并之后，他就做了美大的 CTO。

之前提过，我们团队和道锋团队间做过技术交流，美团一直想做异地多活，但规划了几年一直没做成。后来我问了那边相关同学为什么，他们说道锋虽然是 CTO，但主要管公共技术，也就是搞搞基础设施、大数据之类，业务那边的技术团队非常独立，业务对自己技术团队影响更大。

美团业务的强势度远远超过饿了么，技术要影响业务团队需求比较难，因为如果要做改造，就要有很多技术资源（时间、人员、保障等）投入，一定会影响业务交付，业务不干的。他们业务团队的技术负责人直接汇报给业务老大。

但饿了么不是这样，我进去之后，先是所有的技术都在我这边，但后来发现这样有问题，这个模式不太好，我就跟汪渊商量，把业务的技术跟业务的产品都实线汇报给他，但业务的技术虚线汇报给我。虽然是虚线，但因为汪渊和团队的信任，这个虚线还是比较强的。所以倒不是说罗道锋他搞不定，是业务根本不鸟你，我们这边就比他要好一点，我几乎能完全掌控，即使业务技术团队只是虚线。

分享给需要的人，Ta订阅后你可得 20 元现金奖励

生成海报并分享

赞 9 提建议

© 版权归极客邦科技所有，未经许可不得传播售卖。 页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 13 | 大的决策与难做的决策

下一篇 15 | 兴趣与个人认知

精选留言 (3)

写留言

Condor Hero

2021-11-11

异地多活为什么这么牛，同城灾备的意义在哪，看完文章「搞懂异地多活，看这篇就够了」完全明白了它们的意义，作为用户只能讲，张大神 YYDS。

作者： Kaito
链接： http://kaito-kidd.com/2021/10/15/what-is-the-multi-site-high-availability-design/

10

侯建坡

2021-11-16

文章最后的感触很大，做任何技术的改造，还是要得到业务的支持，我个人认为这种虚虚实实的汇报关系，真是起了很大的作用，没有这层关系，像MT一样，确实有劲儿没处使。

展开

作者回复: 类似软件架构，任何组织架构都有局限性。虚实结合适用饿了么却未必适用美团，更何况，按成王败寇视角，胜利者经验更易被称颂。

2

lesserror

2021-11-16

1. 北京机房（IDC）和 指灾备 IDC，这里的IDC是 互联网数据中心（Internet Data Center）的意思吗？

2. 怎么理解 ：网格化环境呢？

展开

作者回复: 1、是的，一般指物理上聚集在一起的单体数据中心；2、针对外卖场景（打车类似，但简单很多），一般指将地图上大的区域（如：北京五环、上海外环等）划分为适合业务发展或满足用户需要的一个个不规则小区域。同时，针对B、C、D（配送）、BD（销售）等不同角色，区域划分规则也不尽相同（也就是一般说的「多张网」）。例如：B1区域商户服务C1（部分）+C2（部分）用户，同时，D1区域骑手服务B1（部分）+B2（部分）商户，又同时，BD1区域销售服务B3（部分）+B4（部分）商户。

共 2 条评论 >